Displays 34 (2013) 223–240

Contents lists available at SciVerse ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

# Qualitative analysis of individual and composite content factors of stereoscopic 3D video causing visual discomfort

Dong-Wook Kim [a,*], Ji-Sang Yoo [b], Young-Ho Seo [c]

[a] Department of Electronic Materials Engineering, Kwangwoon University, 447-1, Wolgye-1Dong, Nowon-Gu, Seoul 139-701, Republic of Korea
[b] Department of Electronic Engineering, Kwangwoon University, 447-1, Wolgye-1Dong, Nowon-Gu, Seoul 139-701, Republic of Korea
[c] College of Liberal Arts, Kwangwoon University, 447-1, Wolgye-1Dong, Nowon-Gu, Seoul 139-701, Republic of Korea

ARTICLE INFO

*Article history:*
Received 15 February 2013
Received in revised form 4 June 2013
Accepted 3 July 2013
Available online 10 July 2013

*Keywords:*
Viewer's safety
Content production guideline
Discomfort factors
Stereo 3D contents
Disparity

ABSTRACT

Recently, much research has been performed to guarantee viewer's safety in watching stereoscopic 3D videos. Most of them focused on individual content factors causing viewer discomfort. This paper extends the kinds of content factors, and focuses more on the composite, rather than individual, factors. To analyze them, a subjective test for discomfort is performed for four stereoscopic 3D videos. Also, all the quantitative values of the factors to be considered are extracted from the contents. The quantitative factors we consider are the amount of disparity, the frequency and amount of the disparity changes, object movement, and the color and luminance information. In addition, we also include some situational factors, such as story of the contents, situation or circumstances of a scene, and movement, position, and direction of camera. We analyze qualitatively, by comparing the subject test results and the extracted quantitative factors, as well as the situational factors, to find when and how much those factors, in combination or individually, cause viewer discomfort. We summarize, and substantially show which factors, or their combinations, strongly or weakly affect viewer discomfort.



## 1. Introduction

The 3D image industry has been rapidly growing since the movie 'Avatar', because most of the related companies have also launched their prepared products or tools for the acquisition, display, and even editing of 3D video. Therefore, it seems to be almost impossible now to withdraw from the 3D industry. The current 3D images/videos are stereoscopic, based on binocular disparity [1], and are glass-wearing types, where a viewer must wear glasses of active (shutter glasses) or passive (polarized glass) type. Many experts predict that it will progress to multi-view, ultra multi-view, auto-stereoscopic, and finally holographic. But all those technologies, except holography, are also based on binocular disparity, and most of them are of the glass-free type.

There are two problems in the current 3D technology. The first is having to wear glasses, which is cumbersome for viewers. Of course, there are some currently existing glass-free applications, but they are for personal devices, or very restricted areas. The second is discomfort, such as fatigue, dizziness, or even emesis symptom, while watching 3D content. This problem is due to emitting two images compulsively into the corresponding eyes, which causes some difference from the human visual system [2]. Right now, the technologies for glass-free type 3D display seem quite mature, but are not adequate for the second problem, because it is more serious in the glass-free type display. Because the public desire for 3D images is now changing to safety in watching, it must be satisfied by resolving the second problem. Thus, the research on the second problem has been gathering more attention these days, and there have been lots of related works, which will be explained in the next chapter.

The purpose of this paper is to consider as many factors in 3D content as possible, to figure out when, and how strongly, those factors cause viewer discomfort. Here, we focus more on composite factors, rather than the individual factors that most of the existing researches dealt with (some of them considered a few, or a set of factors). For this, we perform a subject test for discomfort, such as dizziness and visual fatigue, to see when and how many viewers feel discomfort. Also, we extract a set of quantitative factors, such as disparity and its change, movement of object, and color and luminance information from the contents for focal point, its vicinity, and the whole image. Also, we consider some situational factors, such as story, situation or circumstance of a scene, position, direction, and movement of camera, which are qualitative. Then, by considering all those factors, and the result from the subject test, we qualitatively analyze the factors and their combinations causing visual discomfort. Finally, we show which factor increases or decreases the 3D effects, with providing empirical data.

* Corresponding author.
*E-mail addresses:* dwkim@kw.ac.kr (D.-W. Kim), jsyoo@kw.ac.kr (J.-S. Yoo), yhseo@kw.ac.kr (Y.-H. Seo).

0141-9382/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.displa.2013.07.002

This paper consists of the following. The next chapter is devoted to explain some related previous works. Then we explain the procedure of our experiments, data extraction, and analyses. In Chapter 4, the results from analyses are explained, which are the basis of the summary and conclusion of this paper in Chapter 5.

## 2. Previous works

We first review the documents announced by international or domestic national institutes. The first one was by the ISO (international standard organization) in 2005 [3]. It included some guidelines related to safety viewing of 3D contents for optical-hypersensitive spasm, dizziness by abnormal visual stress, visual fatigue, etc. The 3D Consortium of Japan has published a document named '3DC Safety Guideline for Popularization of Human-friendly 3D' in 2006 [4], which dealt with some guidelines from the aspect of the viewer watching 3D content. The Telecommunications Technology Association (TTA) [5] of the Republic of Korea has also announced the '3DTV Broadcasting Safety Guideline' in 2010. It consists of four guidelines: 'Guidelines related to viewing environments', 'Guidelines related to viewer', 'Guidelines related to content', and 'Guideline related to display'. Among them, 'Guideline related to display' is the most deeply related to this paper. But it describes in too coarse a manner, such as 'The disparity must not change so rapidly in a very short time' or 'It is desirable to alleviate visual fatigue by smoothly designing zoom or panning'.

The researches related to 3D image/video can be classified as Table 1. The purpose of the researches is either to increase the 3D effect, or to enhance the viewer's safety. In the methodologies, two kinds of approaches have been made, engineering or human factor. That is, there are four categories: engineering technologies to (A) increase 3D effects, (B) to guarantee viewer safety against 3D effects, (C) to increase 3D effects perceptively, emotionally, or psychologically, and (D) to enhance viewer's safety, by reducing visual discomfort, such as visual fatigue and dizziness.

But they are related each other, rather than having strict borderlines. The eventual purpose of all four in Table 1 is to complete a guideline to maximize the 3D effects with maximizing the viewer's safety, and the purpose of this paper also belongs to it. The category of this paper is D, but it is also quite related to C, because it is to find the factors that cause viewer discomfort, but they are highly concerned with excessive 3D effects. Here, we only deal with some results related to human factors, without considering engineering technologies. Because the researches in C and D are inter-related, we focus our explanation more on D.

The recent researches on them have been approached from various areas, such as bioengineering, perception science, psychology, and medical engineering. The representative one on C is [6], which is to find the effects and the human factors affecting 3D, realistic, and immersed feeling. [7] showed to what degree emotional factors, such as recognized functionalities and presence (immersion and involvement) affect 3D feeing. Also, [2] set up the visual attention model for each step during the data processing procedure for 2D or 3D image service, and analyzed the factors increasing attention.

The factor that the researches on D dealt with most frequently was disparity. [8] divided the magnitude of disparity into nine steps, and the degree of defocusing into 5 steps, to show that visual fatigue or discomfort increases as the magnitude of disparity or the degree of defocusing increases. Typically, only the horizontal disparities are considered, but some considered vertical disparities for camera arrangement of parallel type [9], convergence-type [10], or for distortion, such as keystone [11]. Because the current stereoscopy forces the two images into two eyes, the limit of disparity for a human to feel 3D was estimated with a rectangular bar image [12], and random dot stereogram [13]. Also, [14] measured the tolerance of horizontal binocular disparity on a glass-free 3D display.

In Ref. [15] measured the degree of discomfort by SSCQE [16], with respect to the visual accommodation, to show that 3D content makes the difference in the accommodation responses between before, and after, watching. [17] showed that the closer or the farther the object appears than the focal point, the more discomfort a viewer feels. [18] found that asymmetry between the left and the right images, and spatial distortion in combination with disparity, increases the discomfort. [19] targeted a HMD (head-mounted device), but it investigated the matched cases and unmatched cases between the accommodation demand and the vergence demand, to estimate the fusion times for various cases. Those researches support that discomfort, such as visual fatigue in watching 3D content, is caused by visual vergence-accommodation conflict [20], and most of the researchers agree with it. Also, it was medically ascertained with brain waves [21]. Based on the researches above, [22] showed that the causes of visual fatigue are visual stress, pain in the eyes and body, and defocused image. [23] defined the maximum disparity, range of disparity, and object movement as the causes of visual fatigue, and insisted that they could estimate the degree of visual fatigue by measuring them. [24] showed that image complexity according to the number of objects, the amount of texture, and horizontal or vertical movement also cause visual fatigue. There was an attempt to standardize the amount of disparity [25]. There have been several review papers. [26] focused more on the misalignments or excessive imposing, such as stereoscopic distortion and beyond the zone of comfortable viewing and measurement methods. [27] focused more on the content factors, such as disparity distribution, amount of parallax, object variation by motion, and time-varying disparities.

Most of the researches on viewers discomfort so far dealt with disparities, and some considered other factors, such as image distortion, asymmetry between the two images, defocusing, and object movement. However, there could be more factors and their composites that cause more or less discomfort, and it is necessary to find them and their effects, to propose more useful guidelines for 3D content production. We are intending to search all (at least most of) those content factors in this paper. Because the purpose of this paper is to find them in themselves, it deals with them only qualitatively, leaving the quantitative analyses for future work.

**Table 1**
Classification of the researches related to 3D image/video.

| Method | Engineering | Human factor |
|---|---|---|
| 3D effect | A | C |
| Viewer's safety | B | D |

## 3. Experimental method and procedure

This chapter is devoted to developing the method to perform a subject test, the method to extract the content factor, and the procedure to analyze the data from experiment and extraction, which are shown in Fig. 1 as a flow graph. It consists of discomfort estimation, content factors extraction, each content analysis, and synthetic analysis for all the contents. That is, a subject test is performed, to find when and how many subjects feel discomfort for the chosen 3D contents. At the same time, all the quantitative content factors (QNFs) are extracted for the contents. Then the contents are analyzed with the qualitative content factors (QLFs), such as situational factors, as well as the test results, and the extracted QNFs. Each of them is explained below.

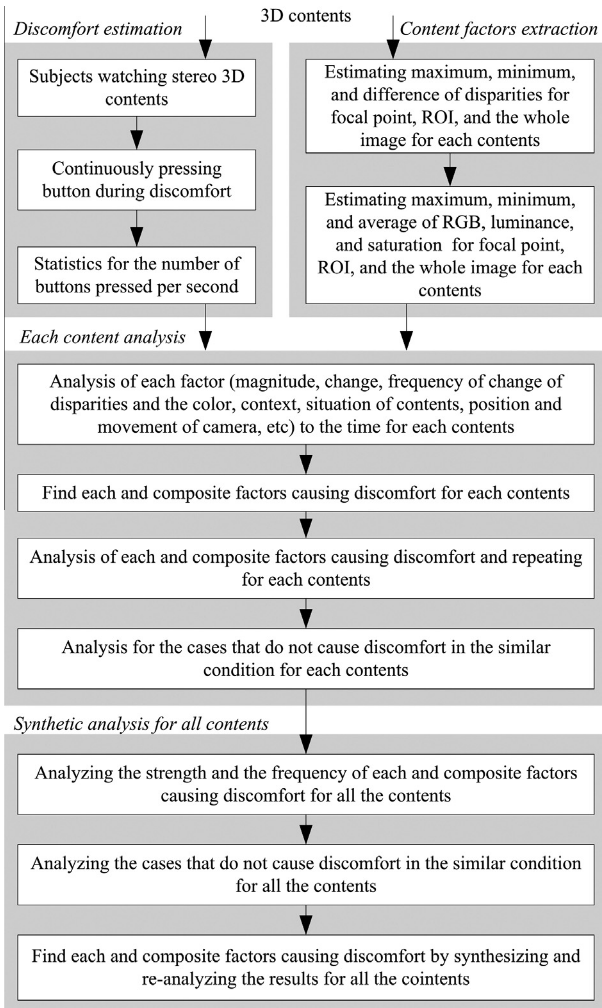


**Fig. 1.** Experimental procedure.

### 3.1. Contents

Four 3D contents [28,29] were used in our experiments, and their information is in Table 2. Two of them are animations, and two are real images. All have HD resolution (1920 × 1080), and their lengths are between about 10 and 18 min. The animations were produced by a 3D content production company, and the real images were by junior or senior university students, whose majors were content production.

'Sun-and-Moon' is a traditional story of the Republic of Korea in which, after a mother, a seller of rice pies, was killed by a tiger, her son and daughter were also chased by the tiger. But they took ropes from heaven, to become the sun and the moon. One peculiar thing is that all the objects and the background look as if they are woven with thick or thin thread, depending on the object. 'An-Dong' is an animated version of a Korean historical war around the 10th century, in which post-Baek-Jae was conquered by the union of Sin-Ra and Korea. There are many war scenes, and many characters appear. Also, it has many unusual scenes in the movement and the position of the camera. 'Ugly Korean' is of two interviewers interviewing two interviewees for new hiring. It implicitly shows the exclusive characteristics of the Korean people. In 'Treasure', two swordsmen fight with swords, and it explicitly give us the message that a treasure cannot belong to a person.

**Table 2**
The information of 3D contents used in the experiments.

| Name | Resolution | Time (sec) | Frames/sec | Class | Producer |
|---|---|---|---|---|---|
| Sun-and-Moon | 1920 × 1080 | 954 | 30 | Animation | Big Eye Inc. |
| An-Dong | 1920 × 1080 | 1079 | 30 | Animation | Big Eye Inc. |
| Ugly Korean | 1920 × 1080 | 703 | 24 | Real Image | Dong-A Inst. Of Media and Arts |
| Treasure | 1920 × 1080 | 755 | 24 | Real image | Dong-A Inst. Of Media and Arts |

**Table 3**
Number of subjects for each content.

| Name | ♯ Of subjects |
|---|---|
| Sun-and-Moon | 31 |
| An-Dong | 25 |
| Ugly Korean | 30 |
| Treasure | 27 |

### 3.2. Subjective test for visual discomfort

We have performed a subjective test to find at which positions of the 3D contents the subjects feel discomfort. The number of subjects who participated in each contents test is listed in Table 3. The test was conducted as follows.

(1) Each subject watched the contents in Table 2, at the distance of 2–2.5 m. We let the subject keystroke a special button continuously, while he or she felt discomfort.
(2) Each subject could watch two contents at maximum. When a subject watched two contents, we let him or her rest for at least 30 min. after watching the first content.
(3) If a subject watched two contents, one should be animation, and the other should be a real image. But in all the cases, the content was chosen randomly.
(4) The number of keystrokes was estimated as the unit of second, by considering the delay from feeling discomfort to stroking the key, such that if a subject stroked the button at least once during a specific time period of second, it was counted as one for that period of second. Thus, the number of those counts for each content became the number of subjects who felt discomfort at that period of a second.

### 3.3. Extracting the values of the quantitative factors

To find the QNFs that might cause discomfort, we have extracted the values of the factors listed in Table 4. We separated each image into focal point, region of interest (ROI), and the rest of the whole image (we will call it just whole image from now on). The focal point was the point that the largest number of viewers gazed at. ROI was arbitrarily determined as 1/16 of the whole image (1/4 of the vertical size and the horizontal size, 480 × 270 in 3D image, 240 × 270 in each of left and right image in the side-by-side format), but it did not include the focal point. We thought the vicinity of the focal point could affect the viewer's feeling. We thought the whole image might not affect the discomfort seriously, but we included it for reference. Of course, the whole image did not include ROI.

**Table 4**
Extracted quantitative content factors.

| Region | Quantitative factors | |
|---|---|---|
| | Kinds | Factors |
| Focal point | Disparity | Disparity |
| | Motion | Amount of motion |
| | RGB | R, G, B values |
| | Luminance/saturation | Luminance, saturation |
| Region of interest (ROI) | Disparity | Maximum, minimum, difference |
| | Motion | Average of the amount of motion<br>Amount of the average motion |
| | RGB | Max., min., average of each of R, G, and B |
| | Luminance/saturation | Max., min., average of luminance and saturation |
| Whole image | Disparity | Maximum, minimum, difference |
| | Motion | Average of the amount of motion<br>Amount of the average motion |
| | RGB | Max., min., average of each of R, G, and B |
| | Luminance/saturation | Max., min., average of luminance and saturation |

For disparity, we assigned negative (positive) when an object looks nearer (farther) than the screen. In estimating disparities, we have used a window-based stereo matching method with variable size of window. The amount of movement was estimated by an optical flow method, with $8 \times 8$ [$pixel^2$] as the size of matching block. For the focal point, only the block including focal point was used in estimating movement. For ROI and the whole image, we estimated Average size of motions and Size of average motion, as the following Eqs. (1) and (2), respectively

$$Average\ size\ of\ motions = \frac{1}{P}\sum_{i=1}^{P}\sqrt{x_i^2 + y_i^2} \tag{1}$$

$$Size\ of\ average\ motions = \frac{1}{P}\sqrt{\left(\sum\nolimits_{i=1}^{P} x_i\right)^2 + \left(\sum\nolimits_{i=1}^{P} y_i\right)^2} \tag{2}$$

where $P$ is the number of $8 \times 8$ blocks, and $(x_i, y_i)$ is the motion vector of the $i$th block. The average size of motions is to focus more on each block, but the size of average motion focuses more on the motion of the whole image. For example, when the motion of each block is different, the average size of motions becomes large, but the size of average motion may be small. But when the camera itself moves, both values become almost the same.

In addition, we extracted the information for color, luminance and saturation. Red (R), green (G), blue (B) are used as the color information. Luminance is the Y component of YUV color format, or I component of HSI format. Saturation is the S component of HSI format [32]. For the focal point, only the absolute values of them were extracted. But for ROI and the whole image, maximum, minimum, and the difference between them for disparity and maximum, minimum, and the average values for color, luminance, and saturation were extracted. Fig. 2 shows an example of those data, in which only the time period of 100–200 s of the ‘An-Dong’ content is included.

## 4. Analysis results

This chapter is devoted to explanation of the results from the analyses. In the explanation, we present evidential data, but because it is too space consuming to show all the data, only the parts directly concerned with each explanation are shown. However in Figs. 3–6, the data for the number of keystrokes and disparities are shown for all the four contents, because the disparity data takes the key role in the analyses. The explanation in this chapter proceeds from global factors to the ones in detail.

### *4.1. Characteristics of contents*

As can be seen from Figs. 3–6, there are many differences in the numbers of keystrokes for the contents. Table 5 shows those numbers, where the number for one person during 1 [min] are also shown, as well as the accumulated number for all the elapsed time, for more objective comparison. In conjunction with the brief explanation for each content in the previous chapter, the characteristics in keystrokes are as follows.

(1) On the whole, more persons key-stroked on the animations, than the real images. This seems to be due to the production property that an animation could change the structure of a scene, amount of disparity, etc. during, or even after production. In addition, the animations in Table 5 were produced by an animation expert company, so that they have many scenes with special effects, which we can see only in a 4D theater in a theme park. But the difference in the number of keystokes of ‘An-Dong’ and ‘Treasure’ is not much. This is because of the image shaking in ‘Treasure’, which will be explained in more detail later.
(2) Dynamic contents have more keystrokes than static ones, as expected. This means that the more dynamic scenes a content has, the more discomfort a viewer feels. This is the same result as previous researches, so we do not mention it further.

### *4.2. Amount of motion*

To compare the amount of motion of an object or image, Table 6 shows the average sizes of motion and sizes of average motions for the focal points, ROIs, and the whole images. The average size of motion of the focal point was the biggest, ROI was next, and the whole image was least. This is because in most scenes, the viewer focuses his or her eyes on the point having the biggest motion. Also, the ROI region is usually connected to the focal point, so that the motion is quite large. A dynamic content had bigger motion than a static one in the animations and the real images, as expected. But one peculiar thing is that the focal point of ‘Sun-and-Moon’ has motion comparable to dynamic content, even though it is static. But its motion in the ROI or the whole image is much smaller than the dynamic ones. From the size of average motion (at the focal point, it is the same as the average size of motion), we can recognize that the cases when the whole image is moving due to the movement of the camera are seldom, or the amounts of movement are small. But ‘Sun-and-Moon’, ‘An-Dong’, and ‘Treasure’ have relatively large whole image movement. In particular, ‘Sun-and-Moon’ and ‘An-Dong’ have quite a lot of image movement, because in an animation, it is relatively easy to make, but for a real image, it may not be.

These are quite similar results to the previous researches [17], that there was a more significant decrease in visual comfort for a scene with large amount of parallax, and large variations in the motion of objects.

### *4.3. Position and movement of camera*

#### *4.3.1. Camera movement*

We can often find the case where the whole image is moving, or the camera itself is moving, and the movement might be small or large, slow or fast. A 2D content having fast movement is dynamic, and is exciting for the viewer, but it does not cause discomfort.

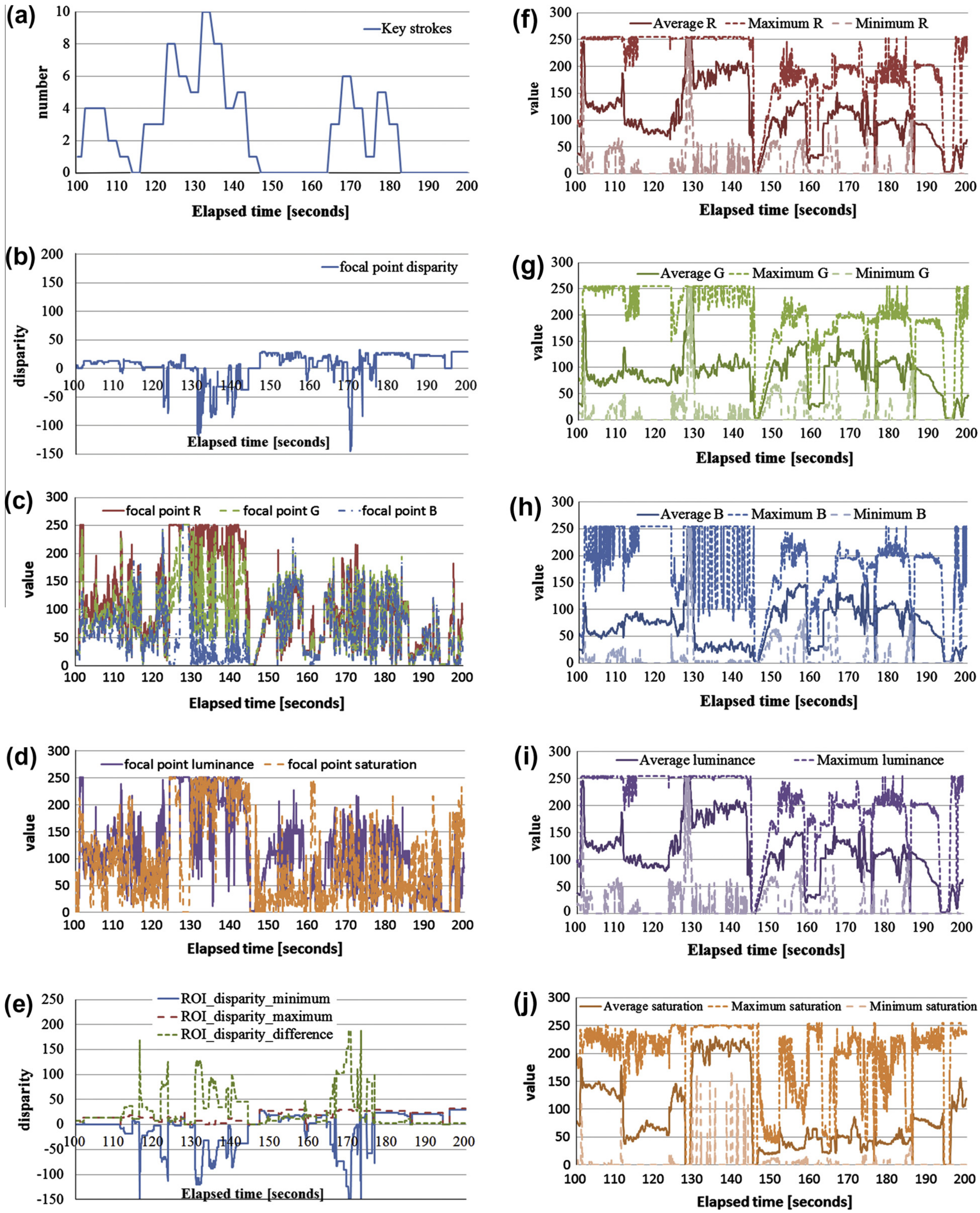


**Fig. 2.** The quantitative data for analysis for the 'An-Dong' content between 100 s and 200 s: (a) number of keystrokes; for the focal point, (b) disparity, (c) RGB, and (d) intensity and saturation; for ROI, (e) disparity, (f) R, (g) G, (h) B, (i) intensity, and (j) saturation; and for the whole image, (k) disparity, (l) R, (m) G, (n) B, (o) intensity, (p) saturation; (q) average size of motion, and (r) Size of average motion.

Here, we are dealing with those cases, to figure out how much they cause viewer discomfort in a 3D content.

*4.3.1.1. Intentional and irregular image shaking or trembling.* There are some cases where the image intentionally and irregularly shakes or trembles. We can see it more often in an animation. For example, when a big object is dropped from a high position in an animation, an image trembling is intentionally inserted with some image effect. Around 710 s of 'Sun-and-Moon' is the scene for the tiger to climb a tree, with chopping the tree with an ax. Whenever he chops the tree, the image is trembled for a while. In this case, some viewers, but not many, felt discomfort. But in

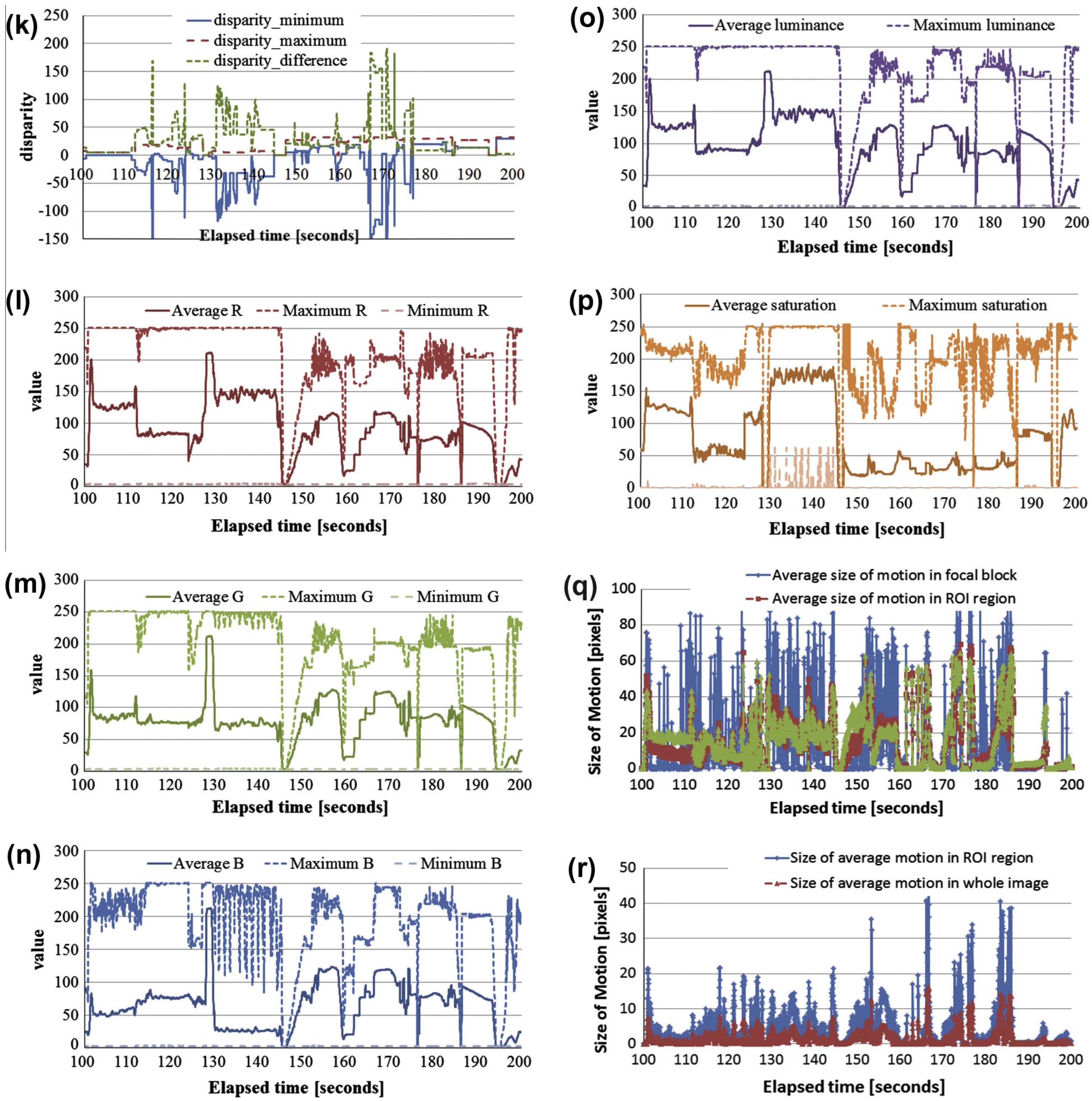


**Fig. 2** (*continued*)

270–300 s and 680–700 s in 'An-Dong' there are several middle and peak keystrokes. These are the scenes of soldiers riding horses. They are shaking, such that the cameraman himself is riding the horse. Those time periods do not have large disparity. Fig. 7 shows (a) an image in this scene, (b) the average sizes of motions, and (c) the sizes of the average motions. From (b) and (c), we can recognize that the image is shaking quite irregularly from 680 s.

Synthetically speaking, for those cases, an intentional and irregular image-shaking or trembling increases the realistic feeling, and also causes viewer discomfort somewhat. Slow but large shaking or trembling causes more discomfort, than small and fast shaking or trembling.

*4.3.1.2. Unintentional and irregular shaking or trembling of image.* Unintentional (that is, due to the environment or situation) image-shaking or trembling appears more in a real image, than an animation. Treasure seems to be captured by two cameramen. Most of the scenes were captured with the camera held by the cameramen, not fixed stably. In particular, the scenes captured by a specific cameraman shakes more than the other, and it is spread all around the content. By considering the keystroke graphs and the disparity graphs in Fig. 6, the scenes in which subjects were not supposed to feel discomfort but some keystroked, even if the numbers are not high, are distributed all around the content, and it seems due to this reason. For example, the keystroke graph starts with some small peaks, and continues to a middle peak and a large one from 120 s to 150 s. This scene shows two swordsmen meeting in a field, with one coming out from woods. It shows two swordsmen alternatively closer and closer. Disparities in this time period are not large nor change a lot, but quite a few viewers felt discomfort. There are many of this kind of situation, such as 25–50 s, around 180 s, 270–300 s, even if their numbers of keystrokes are not large. They have no other reason, except for that.

We can find those situations in other content, too. Around 540 s of 'Ugly Korean' has a middle size of keystrokes, and the data for this part is shown in Fig. 8. This scene shows that one interviewee

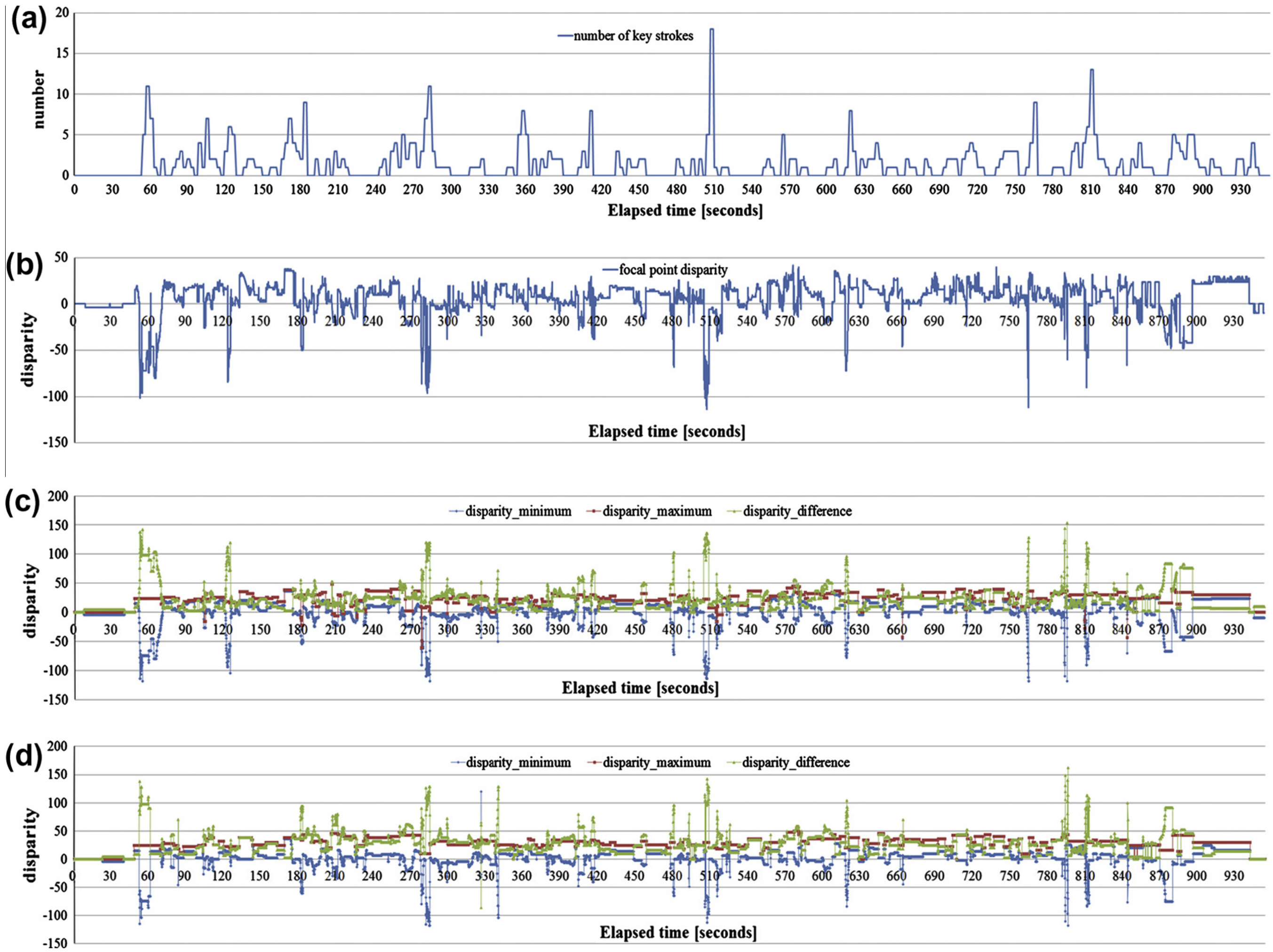


**Fig. 3.** Data for the 'Sun-and-Moon' contents: (a) number of keystrokes; disparities (b) at the focal point, (c) in the ROI, and (d) in the whole image.

is walking along a road after the interview, and the cameraman also seems to walk along with him, holding the camera on his shoulder (Fig. 8(a)). As can see from Fig. 8(b) and (c), two motion values change at the same time from 541 s, but their sizes, durations, or delays are different. That means the whole image is shaking irregularly. Even if the disparity or its change is not large, the number of viewers who felt discomfort was very high.

Those cases look similar to the ones mentioned above of intentional shaking or trembling, but they showed much higher discomfort. This means that in those cases, viewers feel discomfort because of the unstable image, not the disparity.

*4.3.1.3. Fast movement of camera.* It is not hard to see in an animation, or in a real image, that the whole image is moving, due to the movement of the camera. If it is slow enough, it does not cause any discomfort. But a regular fast one-directional or direction-changing image movement causes quite a large amount of discomfort. A window scene from inside a running train corresponds to this case. In an animation, it is relatively easy to produce this kind of scene.

This case appears only in the 'An-Dong' content. An example is around 145 s, and the related data is shown in Fig. 9, in which the camera proceeds rapidly forward through woods. As the camera proceeds, the trees near the camera go backward fast. The motions in Fig. 9(b) and (c) show that the focal point, ROI region, and the whole image have big motions, especially in the ROI region. The number of viewers who felt discomfort here were more than 5. A similar scene is around 680 s, where the number of keystrokes was 4. In this scene, the camera proceeds downward from the top of a valley, then proceeds horizontally through the troops. This scene does not have a big disparity, or its change. So the cause of this discomfort seems to be the fast camera movement.

When the camera moves, the objects around the focal point also appear to move in the reverse direction to the camera. In particular, when the camera moves forward, these objects move backward from the front to the viewer. The factor causing discomfort in this case seems to be a composite one of fast camera movement and the movement of object(s) having large disparity near the focal point, which is explained in more detail later.

*4.3.1.4. Rotating scene.* Sometimes the scene is rotating, due to the camera rotating. It also can be classified into two cases: a fast rotation, in which the object cannot be recognized, and a relatively slow rotation, in which it is possible to recognize the object. A fast rotation often appears when the scene is changed, such as in 'Batman' in the 1960s. This case seldom causes viewer discomfort. But a slow rotation turns out to cause some discomfort, even if not strong. For example the small keystroke peak near 635 s of 'Sun-and-Moon' is by the composite factor that a camera is rotating and lowering with pointing downward (Fig. 10(a)). The disparity is positive, but a few viewers felt discomfort. In the data in Fig. 10(b) and (c), you can see that this scene is rotating and closing-up, from both the motions in the ROI, and the whole image. Those cases

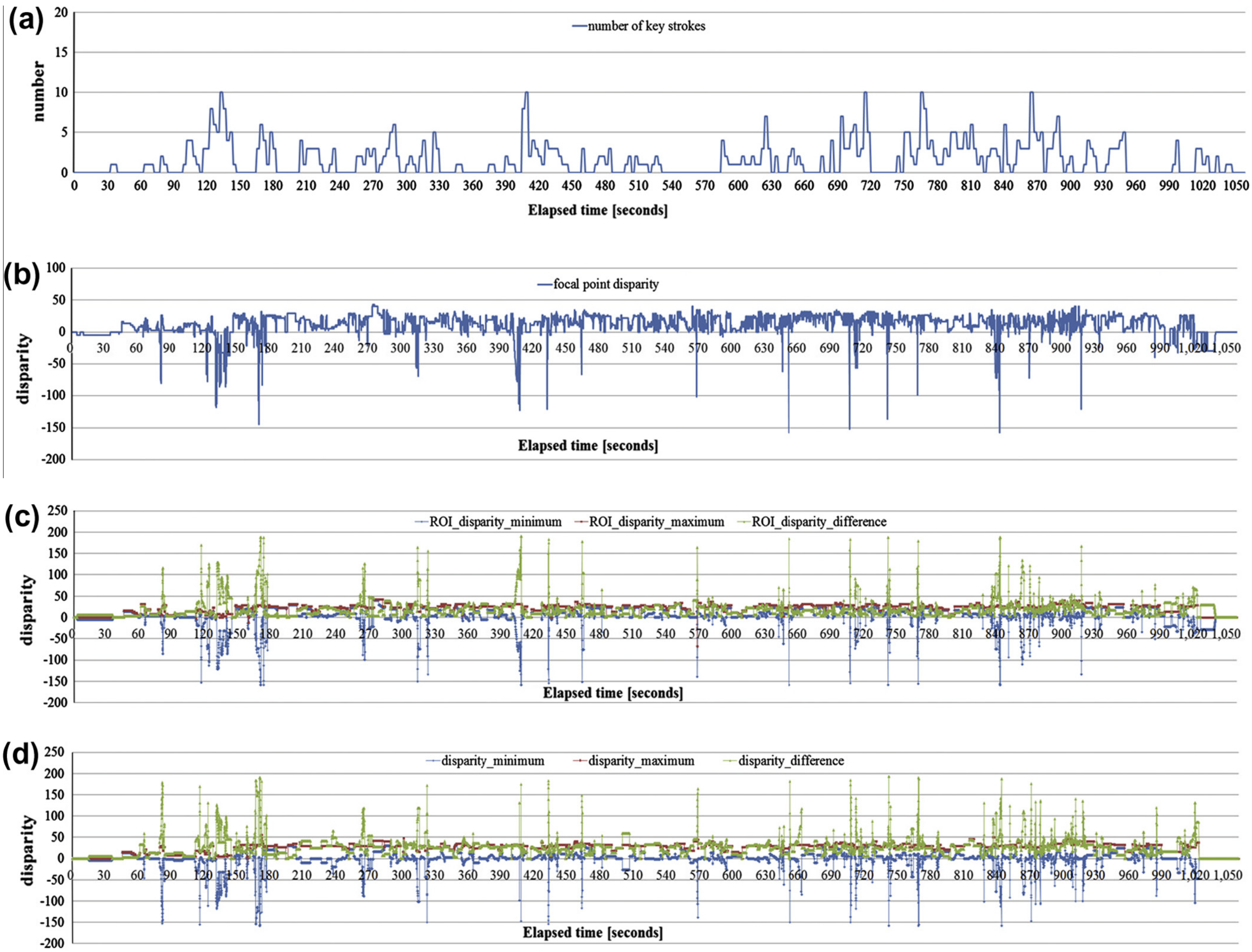


**Fig. 4.** Data for the 'An-Dong' contents: (a) number of keystrokes; disparities (b) at the focal point, (c) in the ROI, and (d) in the whole image.

usually appear when the camera looks downward. That is, camera rotation is usually accompanied with downward camera direction, which is explained next.

#### 4.3.2. Position or direction of the camera

A typical scene is the one captured by a camera at the height of a human, and in the horizontal direction. But sometimes the direction of the camera is downward from a high position, or upward from a low position. In particular, the downward camera can cause some discomfort, similar to the case that we feel dizzy when we look downward from the top of a high building. We can also find this case in the test contents, such as around 370 s of 'Treasure' (Fig. 11(a)) and around 620 s of 'An-Dong' (Fig. 11(b)). The one in (a) is the scene that two swordsmen fight each other, and the one in (b) is that an old woman walks from the enemy camp, gets exhausted, and falls down. Both cases do not have large disparities or motions, but their keystroke numbers are quite large.

The high peak of keystroke in around 400 s of 'An-Dong' has a composite factor. This scene has some horse-riding soldiers from left to right, with capturing upward from a lower position. The soldiers are quite near, so that the disparities are quite large, but not enough for such large keystroke number. So it can be interpreted that the upward camera direction increased discomfort, in addition to the disparities.

These results somewhat coincide with that of the previous works [27,30], in that the normal scene in which the bottom of a scene has more disparity than the top is more comfortable, than such ones explained in this subsection.

### 4.4. Scene change

When a scene changes, the next scene is, in general, unpredictable. But if the story is well known, it may be predictable. Thus, the scene changes by which a viewer feels discomfort can be classified into two kinds.

#### 4.4.1. Unpredictable scene change

Even if the story of a content is well known, there can be some unpredictable scene changes. In such a case quite a few viewers felt discomfort, despite a small disparity. The middle-sized keystroke peak around 360 s of 'Sun-and-Moon' is the scene in Fig. 12(a), which looks like a cartoon. It happens right after the boy hits his younger sister with his finger, while they are playing a game. The flashing star-like object appears twice. The story of this content is very popular in this country, but this scene is totally unpredictable. The disparity in this scene is not that large. Around 940 s of 'An-Dong' (Fig. 12(b)) 3–4 viewers stroked the key. In this scene, the enemy boss is dying. As he dies, a big light beam comes from the sky, and he goes up to the sky through the light beam. Because

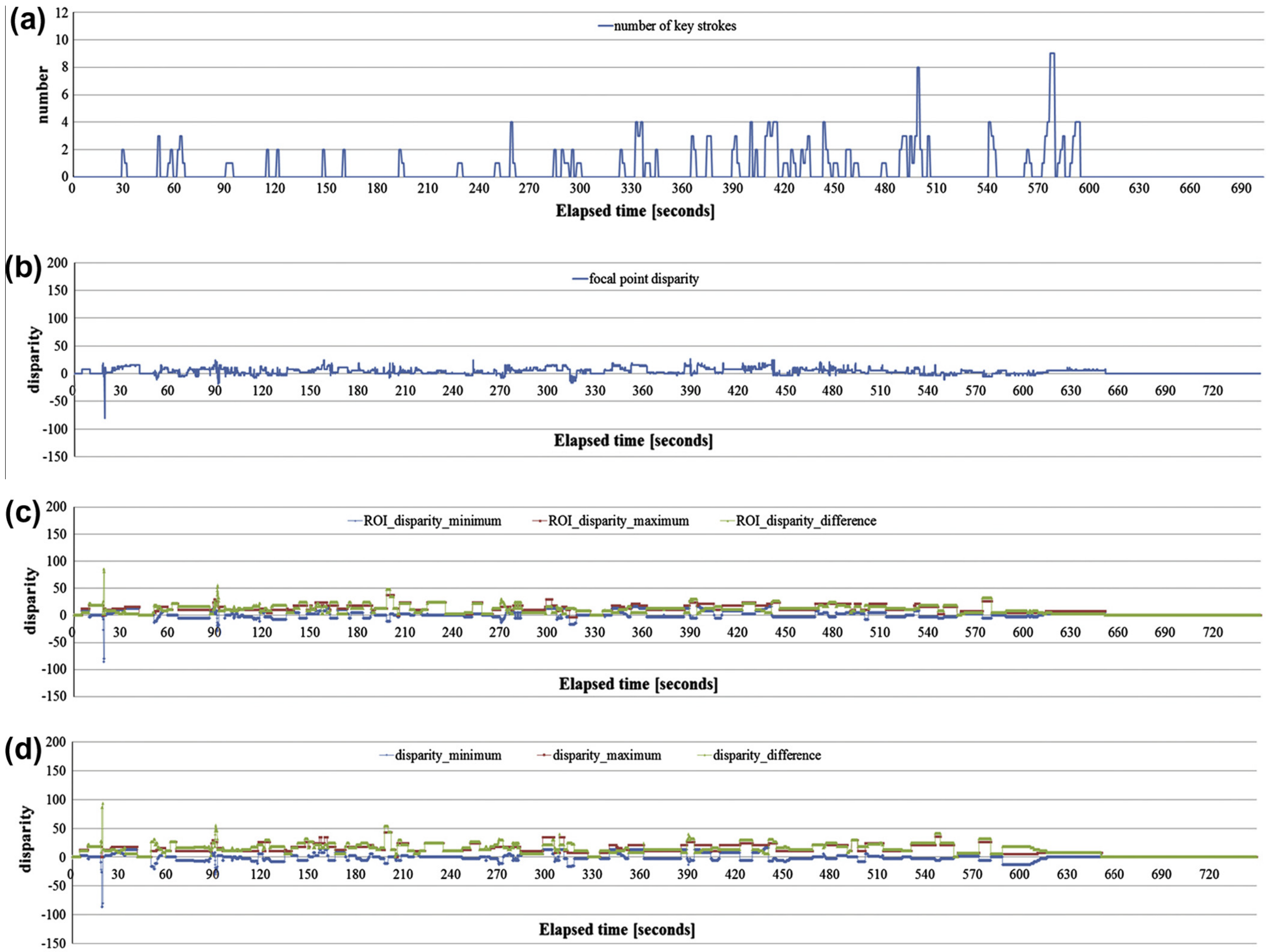


**Fig. 5.** Data for the 'Ugly Korean' contents: (a) number of keystrokes; disparities (b) at the focal point, (c) in the ROI, and (d) in the whole image.

it has a small negative disparity, it should be interpreted as being that an unpredictable scene caused it.

*4.4.2. Repeated or predictable scene*

It is quite inevitable, and appeared many times in the test contents, that a predictable scene decreases discomfort very much, even if its disparity changes to a large negative. An example is the scene around 460 s of 'An-Dong', in which a soldier shoots an arrow, and the arrow comes out to the viewers. But because viewers can predict that the arrow comes to them, only two viewers felt discomfort. Also, around 915 s of 'An-Dong', there is a big negative disparity peak. This is a scene in which the enemy boss and three other side commanders fight while riding horses. Here a sword is heading to the viewers, which has peak negative disparity. But the number of keystrokes is small, compared to the amount of disparity, or its change. In the scene around 510 s of 'Treasure', one swordsman approaches, and finally walks over the camera. The disparity and its change are big enough, but only 2–3 viewers stroked the key.

The cases when the same or similar scenes are repeated also decrease discomfort a lot, even if disparities, or their changes, are large. For example, the scene around 280 s of 'Sun-and-Moon' is of the bad elf floating around the tiger, and this is the first of such scenes. But even though the scenes around 385 s and 490 s. are very similar, the number of keystrokes are notably smaller, that for the first.

*4.5. Disparity*

This section deals with disparity that is known to cause the biggest discomfort. Here, we focus more on the cases that are not recognized by common sense, or are not well known.

*4.5.1. General consideration of disparity*

As can see from Figs. 3–6, large negative disparities are found more in animations, and the disparity changes are also steeper in animations, which seems not unrelated to the numbers of keystrokes in Table 5. From those figures and the table, we can interpret that the discomforts by the two animations are more than the exposed number of keystrokes in Table 5, by considering that both of their stories are well known.

All the contents have positive disparities for most of the time, and only small periods have negative disparities. The case when negative disparity continues for a long time may cause higher discomfort. But the contents chosen do not have any such case that we feel unsatisfactory not to present such data. In general it is clear that the scene with high negative disparity causes high discomfort. But because steep changes of disparities are accompanied in most cases of the test contents, it is hard to say that discomfort certainly depends on the size of disparity, and we will deal with them more later. Also, there were some cases when the numbers of keystrokes are relatively small to their disparities, which will also be mentioned later.

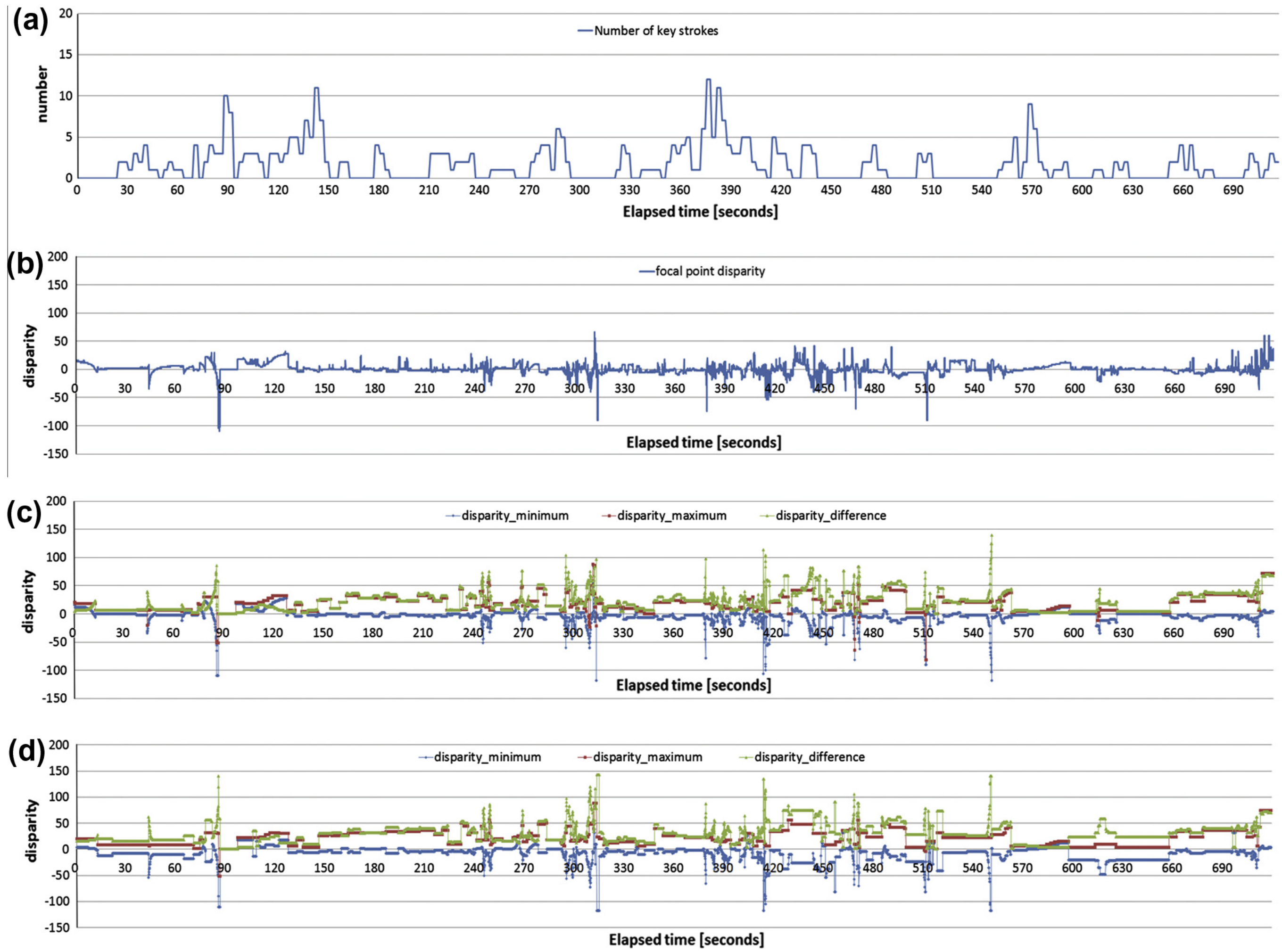


**Fig. 6.** Data for the 'Treasure' contents: (a) number of keystrokes; disparities (b) at the focal point, and (c) in ROI, and (d) in the whole image.

### 4.5.2. Size and change of disparity at focal point

It has been reported that a large negative disparity at the focal point makes a large 3D effect, and causes high discomfort. This subsection deals with disparity, and its change at the focal point.

*4.5.2.1. Sudden unpredictable scene.* The cases when there is an unpredictable object, or one with large disparity at the focal point, can also be divided into two kinds: the case that it is moving and the case that it is not. If it is moving recognizably, it turned out to cause the highest discomfort. An example is the high keystroke peak near 135 s of 'An-Dong', an image of which is shown in Fig. 13(a). In this scene, four Korean letters appear one by one. But each letter comes very close to the viewer, and goes back to its final position. We can find from Fig. 4 that there are frequent large disparity changes in the focal point, ROI, and the whole image. An image of another example is shown in Fig. 13(b). In this scene, an eagle (or hawk) flies to the viewer. But unusually, it comes closer and goes far (as if the camera goes forward and backward), rather than continuously getting closer, which happens three times. Because this scene is much different from the expected, many subjects stroked the key.

We can find this kind of scene in 'Treasure', one of which is near 380 s, where two swordsmen fight. But suddenly the camera closes up on one of them. The disparity and its change are not big around it, but the number of keystrokes is high. It seems to be because this scene is not expected. Also, the keystroke peak near 510 s of 'Sun-and-Moon' can be analyzed in the same way. In this scene, the tiger stretches out his hand through the door (with breaking the door). For this scene the story is that the tiger show his hand to the children by their request, and it is well known; but this sudden action caused ultimate discomfort. There are many other such scenes. This kind of scene may not only maximize the 3D effects, but also cause ultimate viewer discomfort.

**Table 5**
Number of keystrokes for each content.

| Contents | Accumulated ♯ of key-strokes | ♯ Of key-strokes per person per min |
|---|---|---|
| Sun-and-Moon | 1452 | 2.946 |
| An-Dong | 1653 | 3.746 |
| Ugly Korean | 645 | 1.835 |
| Treasure | 1209 | 3.568 |

**Table 6**
Amount of motion for each content.

| Contents | Focal point | Average size of motions | | Size of average motion | |
|---|---|---|---|---|---|
| | | ROI | Whole image | ROI | Whole image |
| Sun-and-Moon | 219.596 | 153.392 | 81.477 | 82.496 | 29.250 |
| An-Dong | 212.897 | 206.489 | 196.544 | 78.827 | 27.949 |
| Ugly Korean | 61.227 | 39.678 | 29.681 | 17.176 | 6.090 |
| Treasure | 237.661 | 202.318 | 172.593 | 60.649 | 16.639 |

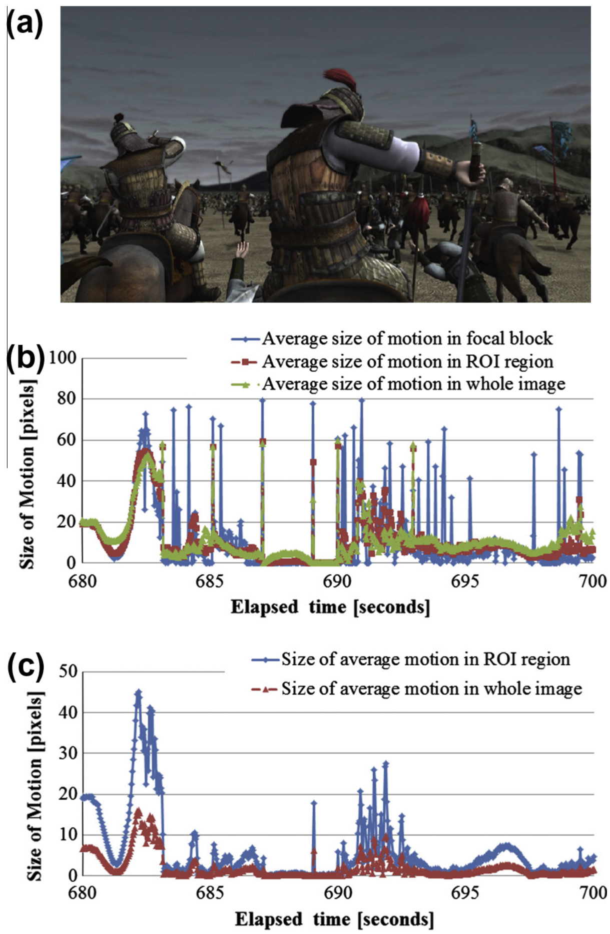


**Fig. 7.** Examples of intentional image shaking (An-Dong): (a) Image at 710 s, (b) average size of motion between 700 s and 720 s, (c) size of average motion between 700 s and 720 s.

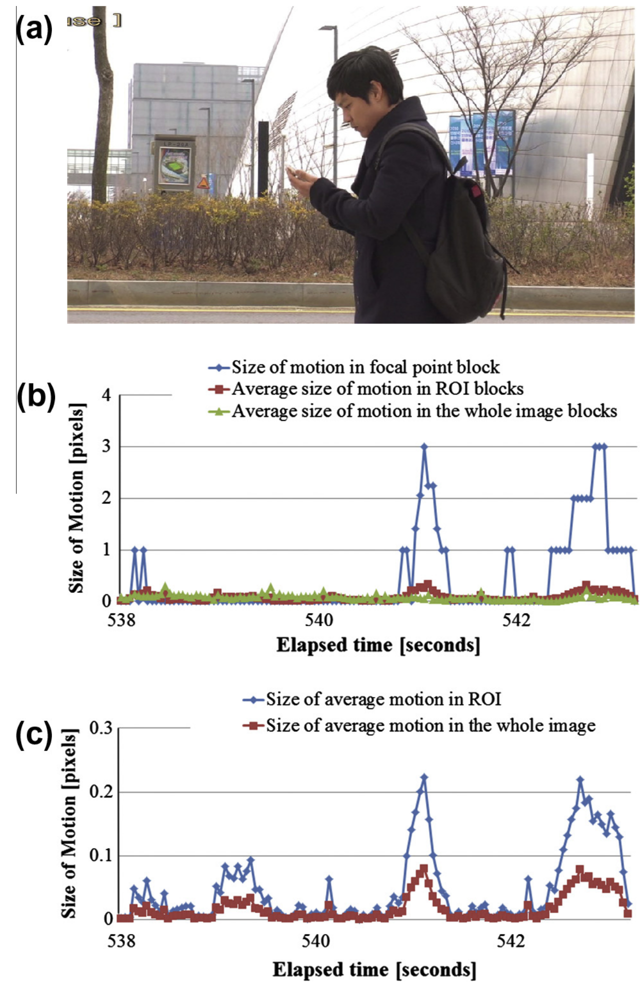


**Fig. 8.** Examples of unintentional image shaking (Treasure): (a) Image at 541 s, (b) average size of motion between 538 s and 543 s, (c) size of average motion between 538 s and 543 s.

*4.5.2.2. Movement of an object with large disparity.* It turns out that there are much difference between moving cases and stationary cases, even if both objects have large disparities. The scene around 315 s of 'Ugly Korean' shows an interviewer talking with pointing to something with his finger. It is a sudden scene change, and a large disparity continues for seconds, but the discomfort in this scene is not large. The scene around 310 s of 'Treasure' shows a swordsman who is very close. In spite of a sudden scene change, and a large disparity continuing for seconds, only one or two viewers felt discomfort. Around 125 s of 'Sun-and-Moon', the tiger suddenly appears in front of the woman, which is shown in Fig. 14(a). This is a sudden scene change, and has large disparity, but the number of viewers who felt discomfort was not large. Those cases have the conditions for viewers to feel large discomfort, such as sudden scene change and large and continuing disparity. But because the objects are still, without moving, the viewer felt much less discomfort to the factors values.

On the other hand, an object not only having a large disparity, but also moving, turns out to cause larger discomfort. The scene around 120 s of 'An-Dong' shows some dead soldiers, and their swords. The tip of one sword is directed to the viewer, and is shaking, according to the sound of distant explosion. Even though it does not have a big disparity, the number of keystrokes was almost the largest. One more example is the scene in which a trolley with some spears is coming out to the viewer, in around 870 s of 'An-Dong', an image of which is shown in Fig. 14(b). When the trolley is coming, it is shaking because of the uneven road, and the spears are shaking too. The disparity around here is not big, but many viewers felt discomfort. There are many other such cases in 'An-Dong' and 'Sun-and-Moon'. That is because they are animations, and are produced by imposing much of this effect.

By considering both cases, we could conclude as follows: it is true that a large-disparity object causes a large 3D effect, but if it is moving, it causes much larger discomfort, as well as a much larger 3D effect. This conclusion partly coincides with the results from previous works [27,31], that the rate of change in disparity magnitude over time causes more discomfort.

*4.5.3. The case when many unspecified objects are coming fast to the viewer*

Those cases appear often in a war scene, for example where many arrows are coming closer from afar. At a glance, it can cause very large discomfort, but there were some differences. That is, when a viewer can fix his (her) eyes on one of them, it causes a big 3D effect and discomfort. But when a viewer cannot, the discomfort is alleviated a lot.

Lets see an example from 'An-Dong'. The scene around 260 s shows many fire arrows coming toward the viewer, which is shown in Fig. 15(a). This time period has many small disparity peaks, and can cause quite a large discomfort, but only 2–3 viewers stroked the key. This kind of scene, of many arrows coming out, also appears around 460 s and the disparity peak is quite big. But this scene also caused discomfort to only two subjects. Also, from

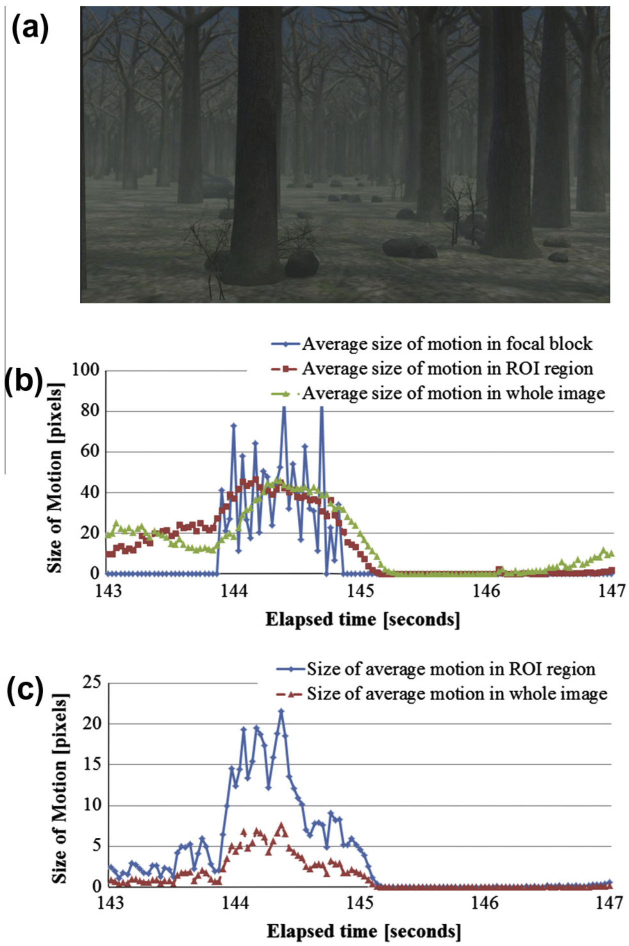


**Fig. 9.** Examples of fast movement of camera (An-Dong): (a) Image at 144 s, (b) average size of motion between 143 s and 147 s, (c) size of average motion between 143 s and 147 s.

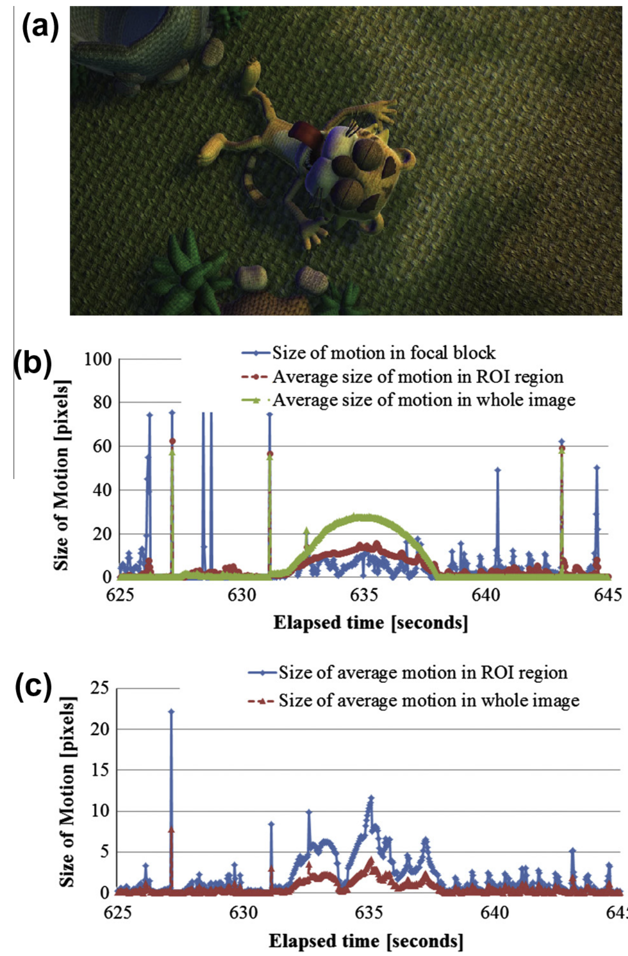


**Fig. 10.** Examples of rotation of the camera (Sun-and-Moon): (a) Image at 645 s, (b) average size of motion between 625 s and 645 s, and (c) size of average motion between 625 s and 645 s.

about 745 s there is a scene in which a water bank is broken by an intentional flood. At first, several water spouts come out toward the viewer, from the openings by cracks. This scene has a large disparity, but only two subjects felt discomfort. On the other hand, the bank itself is broken subsequently, and the broken pieces come out to the viewer, which is shown in Fig. 15(b). The disparity size of this scene is less than the water spout case, but the number of keystrokes was 10.

#### *4.5.4. Moving object(s) in ROI or the whole image*

We can see this kind of scene very often, and it causes quite a large discomfort, even if not ultimate. A representative example is the keystroke peak around 415 s of 'Treasure'. This scene, an image of which is shown in Fig. 16(a), shows a swordsman running through a bamboo forest. Here, because the camera goes along with the swordsman, he is always around the middle of the image, but bamboos pass in the opposite direction. From Fig. 6, we can find larger disparities in the ROI and the whole image, than the focal point.

There is a small keystroke peak around 150 s of 'Ugly Korean'. This is a scene that shows the two interviewers obliquely, but at this time, the far one is speaking, so that the viewers eyes focus on the far one. The disparity on the focal point is positive, but some negative disparities appear in the ROI and the whole image. The number of keystrokes is quite small, because the near interviewer is not moving. But a similar scene at around 500 s has a quite large keystroke peak. This scene shows the two interviewees, but the difference is that the near one is moving. In this scene also, there are only positive disparities at the focal point, but negative disparities in the ROI and the whole image.

There are several similar cases in 'Sun-and-Moon'. The scene at around 170 s is that the tiger suddenly appears in front of the woman, where the woman is shown between the tiger's two legs. Thus, the viewer focuses on the woman. In this scene, the disparity at the focal point is positive, and that of ROI and the whole image is almost 0. These disparities are not enough to cause discomfort, but the size of the keystroke peak was middle. Between 270 s and 300 s, a bad elf appears around the tiger, and attempts to slaughter and eat the woman (Fig. 16(b)). In this scene, the peak of the keystrokes corresponds to the situation that the tiger is speaking, and the elf is floating around in front of the tiger. So, the viewer focuses on the tiger. Here, the disparity in the ROI is larger than that of the focal point.

There are many similar scenes, and most of them turn out to cause quite large discomfort. The larger disparity and the more movement the object in the ROI has, the more discomfort a viewer feels.

#### *4.5.5. High-frequency texture in the background*

As explained before, all the objects in 'Sun-and-Moon' have their surface woven. The thickness of thread is different (refer to Figs. 10, 14(a), or Fig. 16(b)) in different objects, and some can be recognized by the viewer. The scene in 620–645 s shows that the

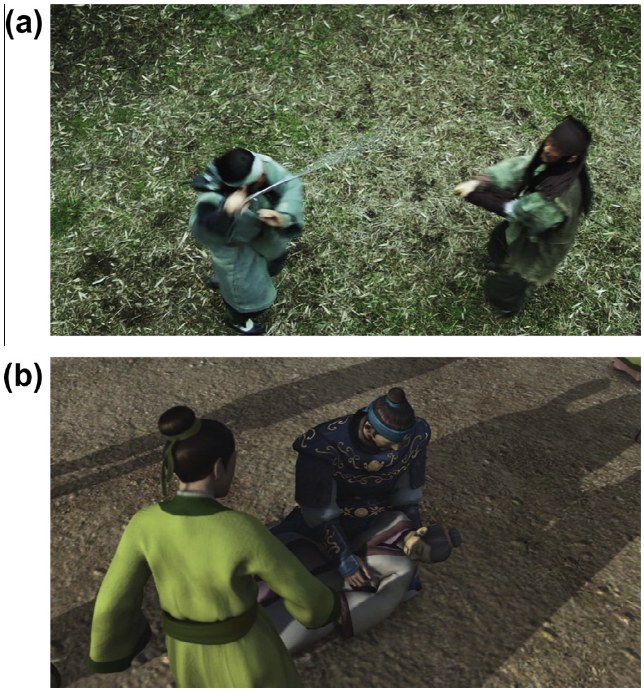

**Fig. 11.** Examples of images captured downward at a high position: (a) near 370 s in 'Treasure' and (b) near 600 s of 'An-Dong'.

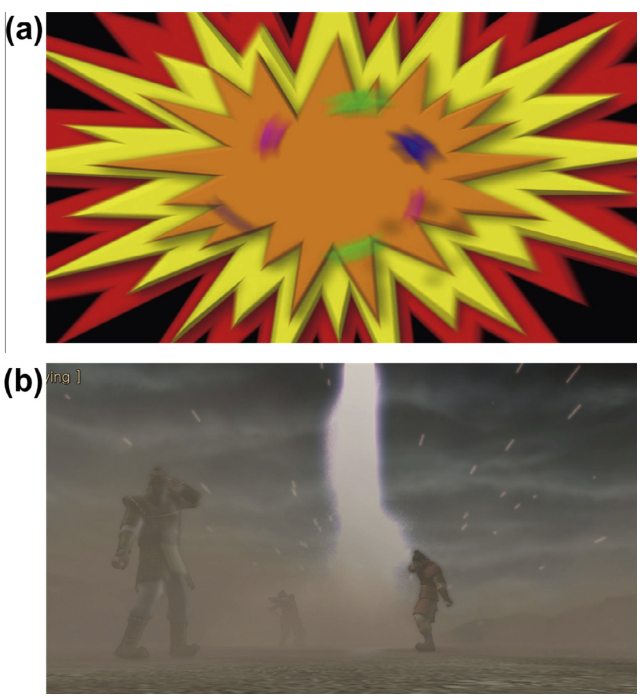

**Fig. 12.** Examples of unpredictable scene: (a) near 360 s in 'Sun-and-Moon' and (b) near 940 s of 'An-Dong'.

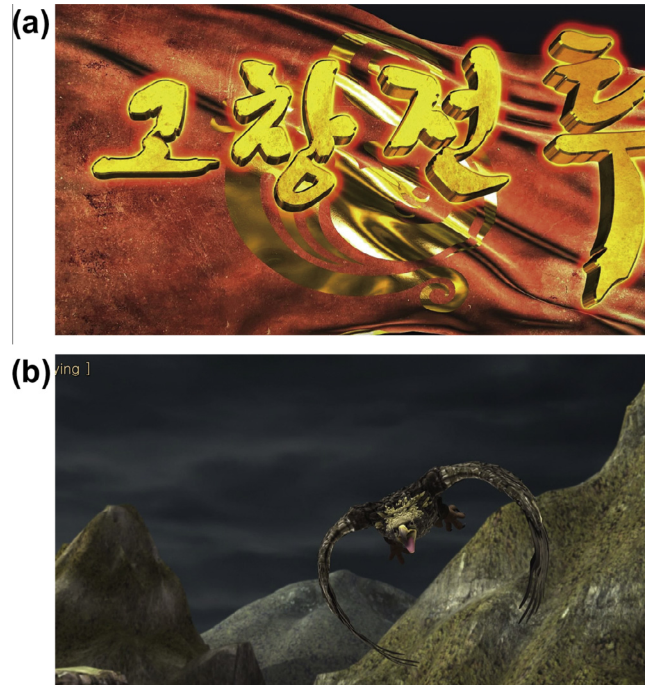

**Fig. 13.** Examples of unpredictable scene: (a) near 135 s in 'An-Dong' and (b) near 400 s of 'An-Dong'.

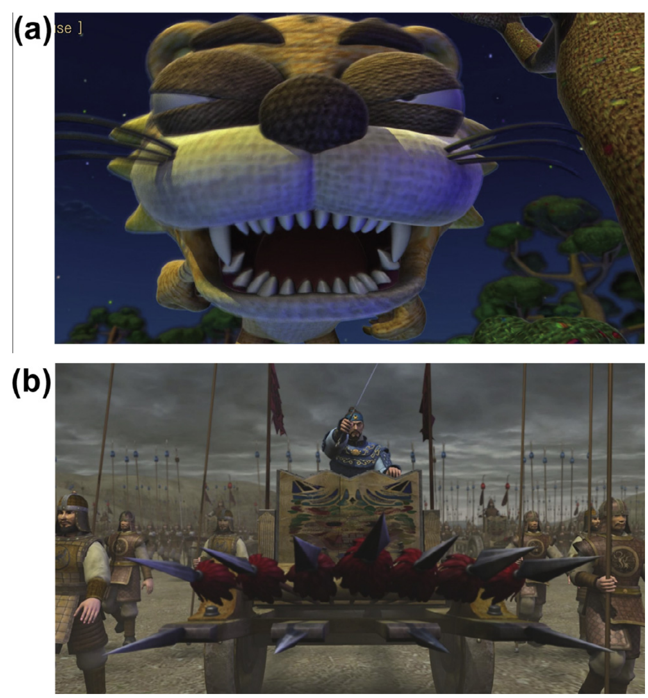

**Fig. 14.** Examples of moving object with large disparity: (a) still object with large disparity (near 125 s in 'Sun-and-Moon') and (b) moving object with large disparity (near 715 s of 'An-Dong').

tiger jumps several times to catch the children on a tree branch, and then falls down on the ground (which is shown in Fig. 10(a)). This scene has been explained as a case of rotating camera, but it is more reasonable that the discomfort is caused by a composite factor of rotating camera and the high-frequency background. Also, around 730–750 s. is the scene that the tiger is climbing the tree while chopping it. It also has been explained before as a shaking image. But because the surface of the tree looks woven with quite thick thread, it can be explained as a composite factor case of shaking image and high frequency background. A noticeable case is around 570 s of 'Treasure', one image of which is shown in Fig. 17. In this scene a swordsman is cut by the opposites sword, and falls down on the ground, which looks like a rough grass field. It was explained as a case of taking a picture downward from a high position. But this scene has a larger keystroke peak than the other similar cases. That means the high frequency

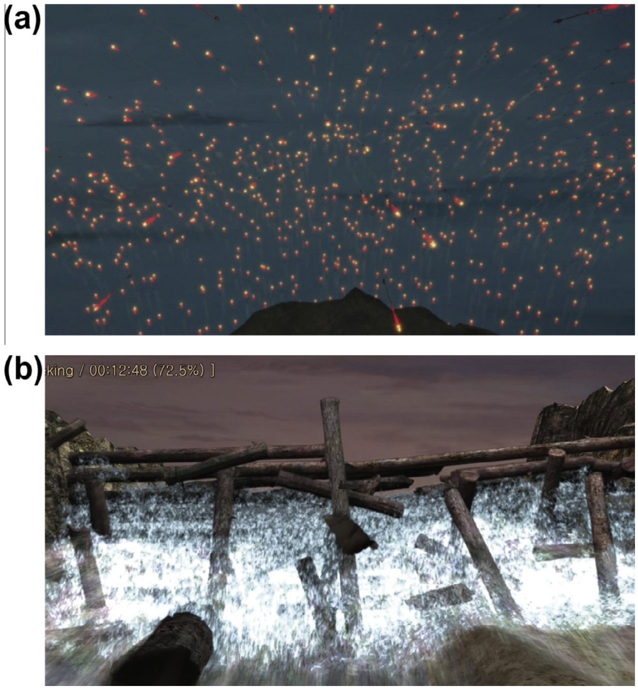


Fig. 15. Examples of multiple objects getting closer: (a) near 260 s in 'An-Dong' and (b) near 765 s of 'An-Dong'.


Fig. 16. Examples of object(s) in ROI closer than focal point: (a) near 415 s in 'Treasure', and (b) near 270 s. of 'Sun-and-Moon'.

background, in addition to the camera position, increases viewers discomfort.

The backgrounds of those scenes consist of high-frequency components, but recognizable textures and the disparities of the object and the background are similar. It makes the viewer experience very confused in the 3D effect and a very unnatural scene, which seems to cause more discomfort than the amount of disparity, or its change.

Fig. 17. Examples of high-frequency background (near 570 s of 'Treasure').

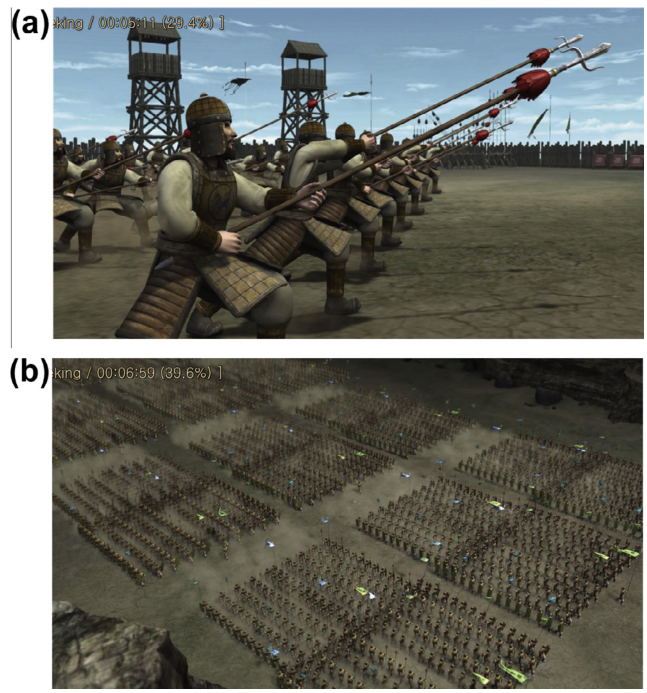


Fig. 18. Examples of many objects that cannot be focused on: (a) near 320 s in 'An-Dong', and (b) near 420 s of 'An-Dong'.

Fig. 19. Example of a scene that is expected to have a large disparity (near 785 s of 'An-Dong').

*4.5.6. Many unspecified objects, but cannot focus on any of them*

Different from the case that many objects come towards the screen, which was explained before, there can be many scenes in which many objects are moving, but they are not coming toward the viewer, or even, a viewer cannot focus on one of them. Those

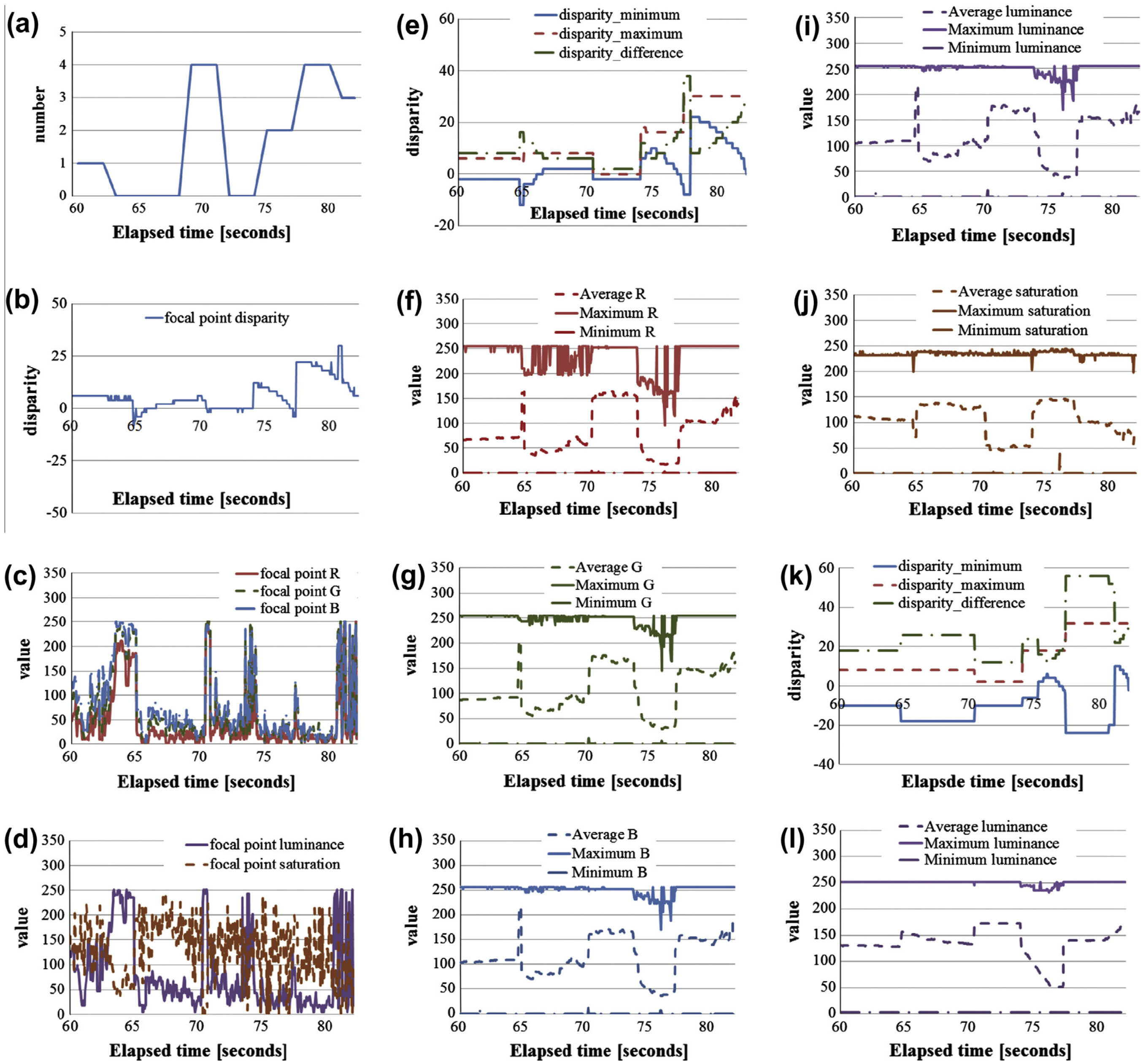


**Fig. 20.** Analysis data during 60–82 s of 'Treasure' contents: (a) number of keystrokes; at the focal point (b) disparity, (c) RGB, and (d) luminance and saturation; at the ROI region (e) disparity, (f) R, (g) G, (h) B, (i) luminance, and (j) saturation; and for the whole image (k) disparity, and (l) luminance.

cases turned out to cause some discomfort. Due to the property of those scenes, many such scenes appear in 'An-Dong'. 320–330 s is the scene in which a troop is training (Fig. 18(a)). 405–420 s shows also a training scene, but it is taken from a high valley far away (Fig. 18(b)), where the formation keeps changing. In both cases, the disparities are always positive everywhere in the image, but some viewers felt discomfort. This is because the viewer cannot focus their eyes on any of the moving objects. In real estimation of the focal point of those scenes, many points were marked, and they did not show any gathering tendency. Therefore, the moving objects that the viewer cannot focus on caused some discomfort.

### *4.5.7. Scene in which a large disparity is expected, but not in reality*

There can be a scene in which a large disparity is expected by circumstances, but in reality the disparity is not large. An example is 780–800 s of 'An-Dong', an image of which is shown in Fig. 19. In this scene, a lot of water flows toward the viewer, after the water bank is corrupted. It is common sense that this kind of scene has a large negative disparity, but in reality most of the scene has positive disparities, and there are only one or two small negative peaks. But five subjects felt discomfort in this scene.

The scene around 360 s of 'Sun-and-Moon' shows the scene in which the boy hits his younger sisters head with his finger, then a large star appears, as in a cartoon (Fig. 12(a)). It was explained before as an unexpected scene. But the star in this circumstance usually has a large disparity. But in reality, the amount of disparity or its change is not noticeable. Thus, it is more reasonable to interpret it as a composite factor of unexpected scene and circumstantially large disparity causing such a big discomfort.

In most other such scenes, quite a few viewers felt discomfort. It seems that many viewers expect a large disparity in those scenes, because the content is 3D video.

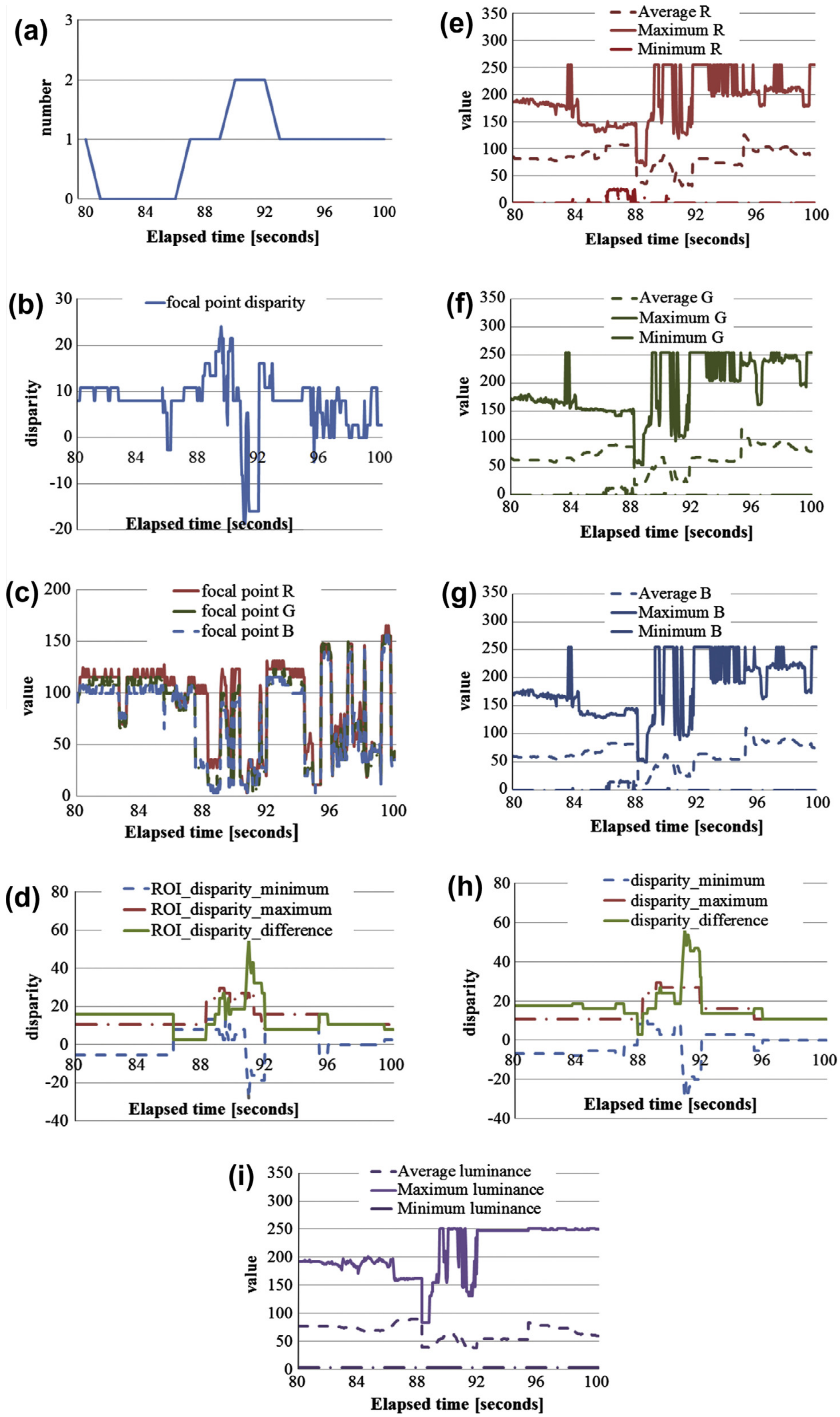


**Fig. 21.** Analysis data between 80 and 100 s of 'Ugly Korean' contents: (a) number of keystrokes; at the focal point (b) disparity, and (c) RGB; at the ROI region (d) disparity, (e) R, (f) G, and (g) B; and for the whole image (h) disparity, and (i) luminance and saturation.

### *4.6. Effect of color factors*

A few years ago, a Japanese child fainted while he was watching a pocket monster cartoon, and it has been a hot issue for a while. The fact that this happened, even though the cartoon was 2D, not 3D, implicitly shows that color factors can cause visual fatigue, or even discomfort. Thus, this section is to deal with color factors connected to the 3D effect. But their effects are not larger than the disparity factors, and it is not easy to closely examine them, which is because they act in composite with other 3D effects.

**Table 7**
Summary of the factors explained as causing discomfort.

| Individual or composite factors | | | Strength of discomfort |
|---|---|---|---|
| Classes | Factors | Cases | |
| Content characteristics | Animation is stronger than real image.<br>Dynamic content is stronger than static one. | | |
| Movement | Stronger in content with larger movement | | |
| Position and movement of camera | Movement of camera | Intentional and irregular shaking or trembling image | Middle |
| | | Unintentional and irregular shaking or trembling image | Middle |
| | | Fast camera movement | Strong |
| | | Camera rotation | Weak |
| | Downward or upward shooting | | Middle |
| Scene change | Unpredictable | | Strong |
| | Repeated or predictable | | Weak |
| Disparity | ROI | stronger when faster disparity change is accompanied | |
| | Focal point | Unpredictable large-disparity object is moving | Strong |
| | | Large-disparity object is moving | Strong |
| | | Large-disparity object is not moving | Weak |
| | Many objects are approaching the viewer fast, but the viewer cannot focus on any of them | | Weak |
| | Many objects are approaching the viewer fast, and the viewer can recognize any one of them | | Strong |
| | High-frequency components in background | | Strong |
| | Many unspecified objects so that a viewer cannot focus v | | Middle |
| | A scene that is predicted to have large disparity according to the scene situation, but has smaller in reality | | Middle |
| | Tiled shooting for several objects with almost the same size, but focus on the nearer ones | | Weak |
| Color or luminance | Large and sudden change in color or luminance | | Increase |
| | Low luminance | | Decrease |

Therefore, this section deals only with the cases that clearly appear.

It turns out that color factors increase or decrease discomfort, as well as the 3D effect. The case of increasing discomfort appears in a real image. Thats because it is not easy to adjust color factors in a real image, once it is captured. The case of decreasing discomfort appears in both animation and real images. The two cases are explained with some examples.

### *4.6.1. The case when color factors cause or increase discomfort*

The case when color factors cause discomfort appears sometimes. The representative example is around 70 s of 'Treasure', where four subjects felt discomfort. In this scene, one of the swordsmen runs away from the camera. The data for analysis is shown in Fig. 20. Here, only luminance and saturation are shown for the whole image. First, disparities and their changes at the focal point, ROI, and the whole image are not large. But the RGB values at the focal point show steep peaks at around 70 s, and the luminance does too. The RGB values in the ROI change frequently, and the average values increase steeply around 70 s (the luminance is similar). Also, any situational factor does not seem to cause any effect. Thus, it should be interpreted that the steep increase in color factors caused the discomfort.

The key strokes around 180 s of 'Treasure' cannot be caused by disparity, because all the disparities at the focal point, ROI, and the whole image are almost 0, and do not change. This scene alternatively shows the two swordsmen one after another, where one is standing in the shadow, and the other in the sunlight. So the luminance, as well as the RGB changes suddenly, whenever it changes the swordsman. They are standing still, the situational factors cannot be affective, but four subjects stroked the key.

Those cases should be interpreted as the sudden change in color or luminance increasing the effect of disparity, so that viewer may feel discomfort, more than the amount of other factors.

### *4.6.2. The case when color factors decreases discomfort*

There appear some cases when the color factors decrease discomfort. For example, let's look at the data around 90 s of 'Ugly Korean'. The data is shown in Fig. 21, where only luminance and saturation are shown for the whole image. From Fig. 21(b), we can see that big disparity changes start from around 90 s, and continue to 92 s. But the number of keystrokes was only two, which is too small as a disparity factor. This scene shows that the two interviewers come to the interview room, but only their feet are shown. As in Fig. 21(c), RGB values at the focal point during 88–92 s are very low, and the luminance is, too. Also, the average values of RGB and luminance in the ROI or the whole image remain low, as in Fig. 21(d)–(f) and (i).

There are some similar scenes in other test contents. There is a big disparity around 510 s of 'Treasure', and it is a scene in which a swordsman runs to the camera, and finally over the camera. Here, only 2–3 subjects stroked the key, which is very small, compared to the size of the disparity. This scene also has so low RGB or luminance values that discomfort is decreased, we think. Also, only two subjects felt discomfort around 480 s of 'Sun-and-Moon', even though there are large disparities and their changes. This scene shows that the tiger spreads flour on all his body, to fool the children. Here, the elf scatters the flour all around, so that the whole scene becomes gray. This scene is also affected by the color factors.

## 5. Summary and conclusion

In this paper, we have examined qualitatively the individual and composite content factors causing viewer discomfort. As the method, we have performed subject tests, with the test contents consisting of animations and real images. At the same time, we have extracted all the quantitative values of the content factors that might cause discomfort. Then the results from the subject test and the extracted factors were used to analyze the contents, with considering some qualitative factors, such as story of content, camera position and movement, and environmental circumstances.

The analysis results are summarized in Table 7, with the strength of each factor or case. Disparity turned out to be the most important factor that causes discomfort, as the results from previous works. But the largest discomfort was caused when it acted in composite with other factors, rather than alone, as simply a large disparity. That is, the case when a large disparity is accompanied with fast disparity changes caused the biggest discomfort. Also, a moving large-disparity object caused much larger discomfort than a still one. When a large-disparity object was moving in the ROI, many subjects felt discomfort. When many unspecified objects were coming to the viewer, the strength of discomfort was much different, according to the recognition of the viewer. That is, the case when the viewer could focus eyes on each of the objects caused much higher discomfort, than the case when not. Also, when a large-disparity object is moving, the case that the movement is predictable causes much lower discomfort, than the unpredictable case.

An irregular shaking or trembling scene and abnormal camera position, movement, or moving direction turned out to be causes of discomfort. High-frequency background around the main object might cause big discomfort. Also a situation that makes a large disparity, but not in reality, could cause quite a large discomfort. But when similar scenes are repeated, the discomfort was much reduced. The sudden and frequent change in color and luminance might cause discomfort. But low luminance might reduce discomfort.

From the analysis results, we can conclude that composite factors, rather than a single factor such as disparity, cause bigger discomfort. Also, big disparity does not always cause a big discomfort, and discomfort is highly sensitive to disparity change, and its frequency. In addition, story content, situation, camera movement and direction, even color factors, as well as disparity, affect discomfort, and their composites cause much larger 3D effects and discomfort.

This paper has paid more attention to qualitative analysis of the factors. But to make a set of reasonable guidelines for content production, the factors and their composites need to be analyzed quantitatively, to get a set of values for a content to keep, such as strength and duration time. That is, it is necessary to measure experimentally how much discomfort it causes, for how strong, and how long, a factor or a composite continues. For this, it is necessary that for a factor or a composite, a content that can be used to measure each or some of the factors should be produced, a more accurate experimental method should be architected, and more objective analysis should be performed. Also, those kinds of research must be done by many associated areas' cooperation or collaboration, which is also our expectation.

## 6. Acknowledgement

This work (2010-0012898) was supported by Mid-career Researcher Program through NRF grant funded by the MEST.

## References

[1] O. Schreer, P. Kauff, T. Sikora (Eds.), 3D Video Communications, John Wiley & Sons Ltd., The Atrium, England, 2005.
[2] Q. Huynu-Thu, M. Barkowsky, P.L. Callet, The importance of visual attention in improving the 3D-TV viewing experience: overview and new perspectives, IEEE Transactions on Broadcasting 57-1 (2011) 421–431.
[3] IWA3: ISO, International Workshop Agreement 3, 2005.
[4] 3D Consortium, 3DC Safety Guideline for Popularization of Human-friendly 3D, 2006.
[5] Telecommunications Technology Association (TTA), 3DTV Broadcasting Safety Guideline TTAK.KO-07.0086, December 23rd, 2010.
[6] T. Inoue, K. Uehira, K. Komiya, M. Suzuki, Human factors of 3-D displays for visual reality, in: IEEE Industrial Application Society Annual Meeting 2008, pp. 1–4.
[7] D.H. Chung, H.C. Yang, Exploratory analysis of 3D stereoscopic video measurement, Journal of the KSDIM 6-4 (2010) 225–235.
[8] Matthias WÅpking, Viewing comfort with stereoscopic pictures: an experimental study on the subjective effects of disparity magnitude and depth of focus, Journal of the Society for Information Display 3/3 (1995) 101–103.
[9] L. Stelach, W. Tam, F. Speranza, R. Renaud, T. Martin, Improving the visual comfort of stereoscopic images, Proceedings of SPIE 5006 (2003) 269–282.
[10] W. Ijsselsteijn, H. Ridder, J. Vliegen, Subjective evaluation of stereoscopic images: effects of camera parameters and display duration, IEEE Transactions on CSVT 10-2 (2000) 145–154.
[11] R. Woods, T. Docherty, R. Koch, Image distortion in stereoscopic video systems, Proceedings of SPIE 1915 (1993) 36–48.
[12] ITU, Subjective Assessment of Stereoscopic Television Pictures, ITU Recommendation BT. 1438, 2000.
[13] C. Erkelens, Fusional limit for a large random-dot stereogram, Vision Research 28 (1998) 345–353.
[14] T. Iwasaki, T. Kubota, A. Tawara, The tolerance range of binocular disparity on a 3D display based on the physiological characteristics of ocular accommodation, Display 30 (2009) 44–48.
[15] S. Yano, S. Ide, T. Mitsuhashi, H. Thwaites, A study of visual fatigue and visual comfort for 3D HDTV/HDTV images, Displays 23 (2002) 191–201.
[16] T. Alpert, J. Evain, Subjective quality evaluation the SSCQE and DSCQE methodologies, EBU Technical Review (1997) 12–20.
[17] S. Yano, M. Emoto, T. Mitsuhashi, Two factors in visual fatigue caused by stereoscopic HDTV images, Displays 25 (2004) 141–150.
[18] F. Kooi, A. Toet, Visual comfort of binocular and 3D displays, Displays 25 (2004) 99–108.
[19] R.H.Y. So, W.S. Wong, R. Yip, A. Lam, P. Ting, Benefits of matching accommodative demands to vergence demands in a binocular head-mounted display: a study on stereo fusion times, Presence: Teleoperators & Virtual Environments 20 (6) (2011) 545–558.
[20] D. Hoffman, A. Girshick, A. Akeley, M. Banks, Vergence-accommodation conflicts hinder visual performance and cause visual fatigue, Journal of Vision 8:33 (3) (2008) 1–30.
[21] H. Li, J. Seo, K. Kham, S. Lee, Measurement of 3D visual fatigue using event-related potential (ERP): 3D oddball paradigm, in: IEEE 3DTV-Con08, 2008.
[22] H.C. Li, Human factor research on the measurement of subjective three dimensional fatigue, Journal of Broadcast Engineering 15-5 (2010) 607–616.
[23] D.H. Kim, S.H. Choi, K.H. Sohn, Visual fatigue prediction for stereoscopic video considering individual fusional characteristics, Journal of Broadcast Engineering 16-2 (2011) 331–338.
[24] N.K. Kim, Human factors for 3D stereoscopic images, Journal of Information Scientists and Engineering 29-8 (2011) 48–54.
[25] M. Lambooij, W.A. Ijsselsteijn, I. Heynderickx, Visual discomfort of 3D TV: assessment methods and modeling, Displays 32 (2011) 209–218.
[26] M. Lambooij, M. Fortuin, I. Heynderickx, W.A. IJsselsteijn, Visual discomfort and visual fatigue of stereoscopic displays: a review, Journal of Imaging Science and Technology 53 (3) (2009) 030201-1–030201-14.
[27] W.J. Tam, F. Speranza, S. Yano, K. Shimono, H. Ono, Stereoscopic 3D-TV: visual comfort, IEEE Transactions on Broadcasting 57 (2) (2011) 335–346.
[28] http://bigient.com/en/sub/work_01_02.php.
[29] http://3d.dima.ac.kr/3Dgallery/index.asp?category=1.
[30] Y. Nojiri, H. Yamanoue, S. Ide, S. Yano, F. Okana, Parallex distribution and visual comfort on stereoscopic HDTTV, Proceedings of the IBC (2006) 373–380.
[31] F. Speranza, W.J. Tam, R. Renaud, N. Hur, Effect of disparity and motion on visual comfort of stereoscopic images, Proceedings of the Stereoscopic Display and Visual Reality System, xiii 6055 (2006). 60550B-1-9.
[32] R.C. Gonzalez, R.E. Woods, Digital Image Processing, third ed., Prentice-Hall, Upper Saddle River, NJ, 2008.